Mock data sensitivity studies

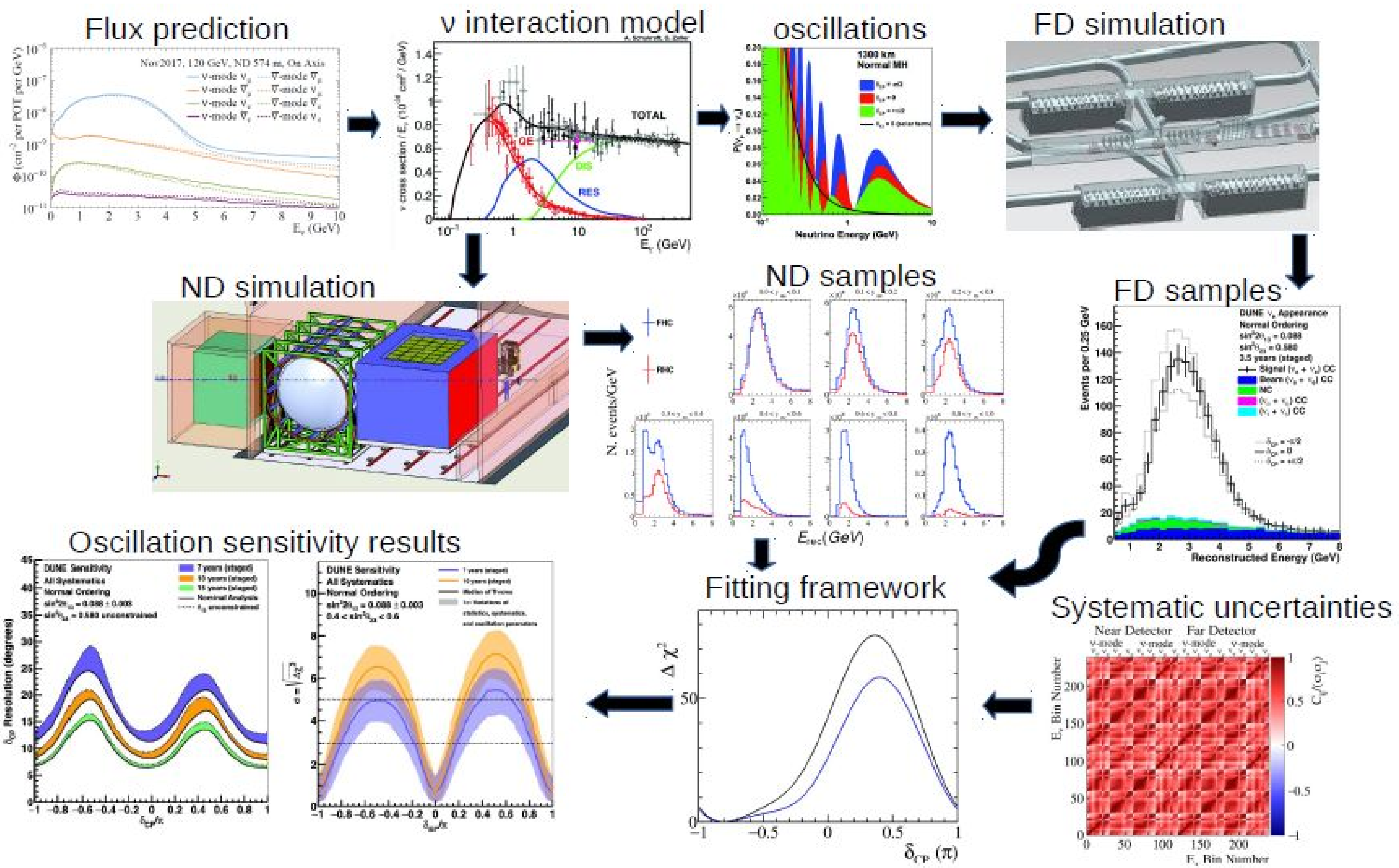
Chris Marshall
Lawrence Berkeley National Laboratory
2 October, 2019

DUNE

Outline

- LBL sensitivity analysis with ND: overview
- Baked-in assumptions & limitations
- Why this is not scalable to ND TDR
- Basic idea of mock data
- Advantages of mock data over full-blown sensitivity

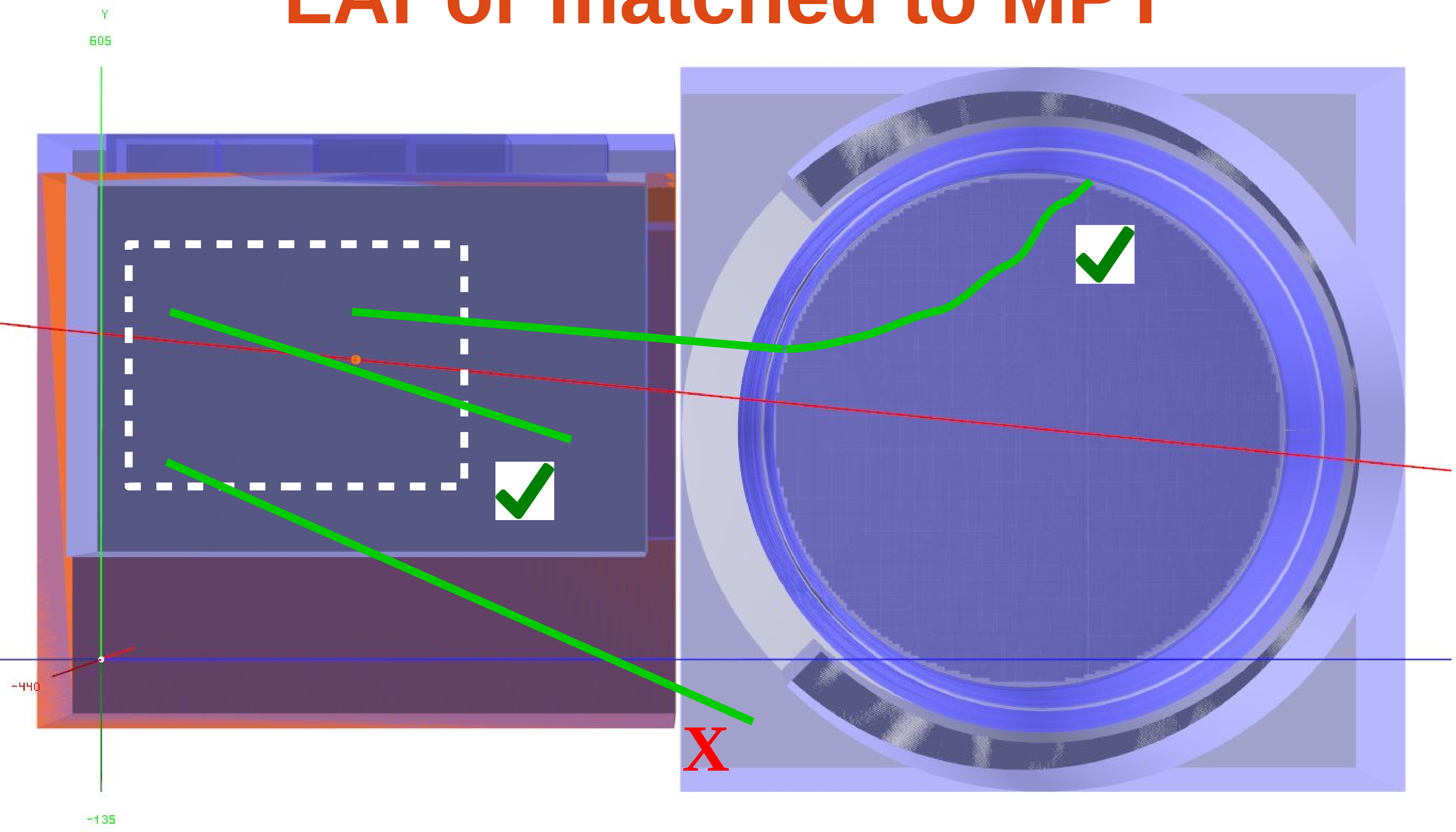
LBL sensitivity analysis



Near detector strategy for FD TDR analysis

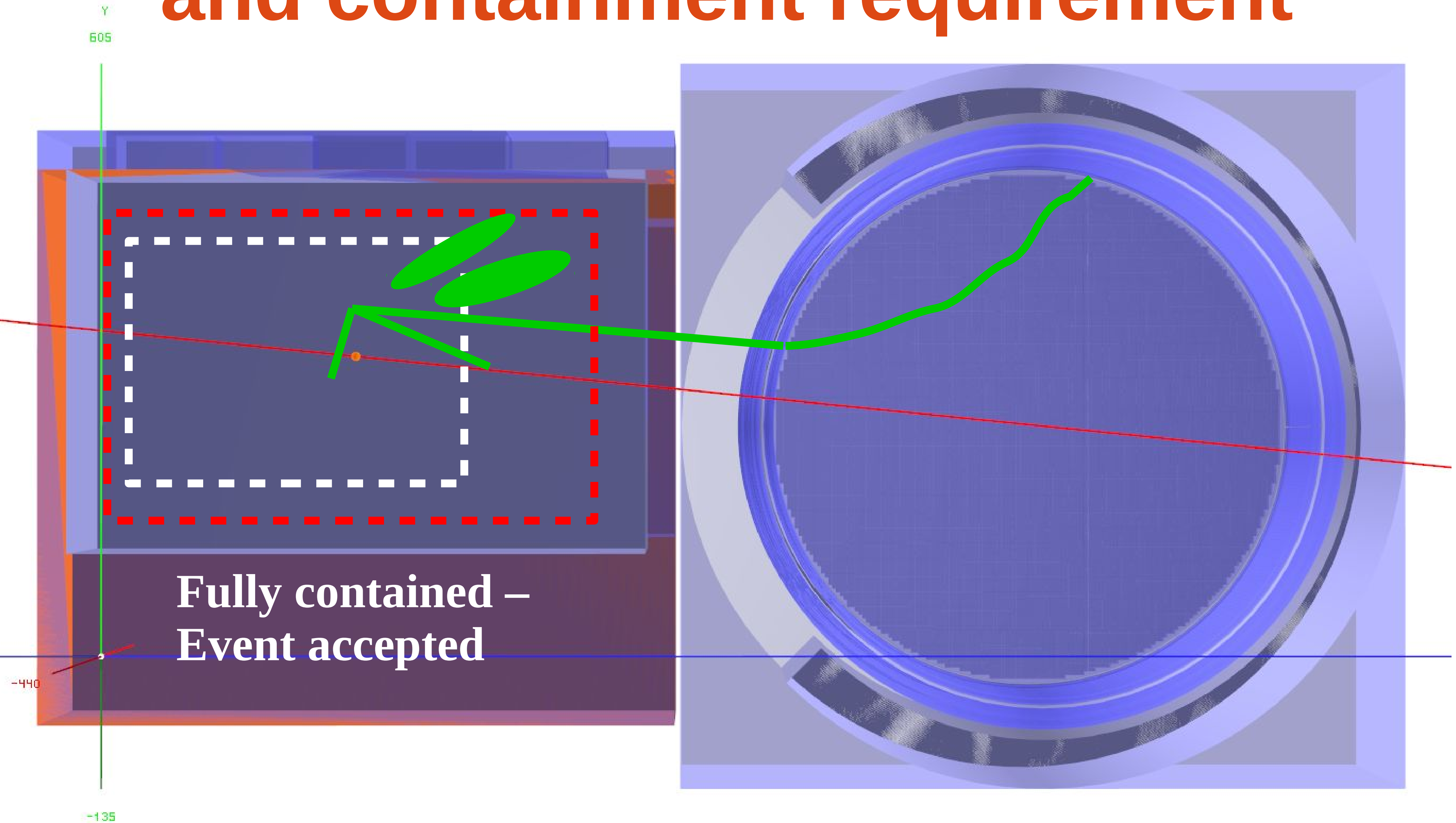
- Generate events with GENIE, with same reweighting framework used in FD
- Propagate final-state particles through detector geometry with Geant4
- Pseudo-reconstruction based in Geant4 energy deposits
- Form LAr ν_μ CC samples to use in analysis

Muon acceptance: contained in LAr or matched to MPT

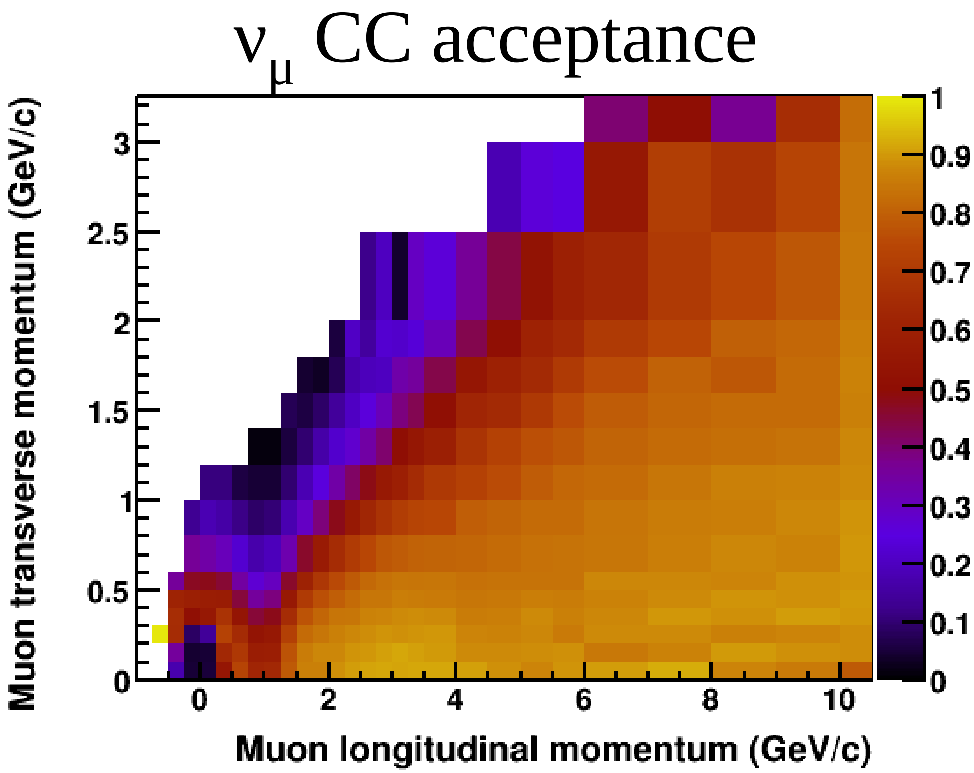


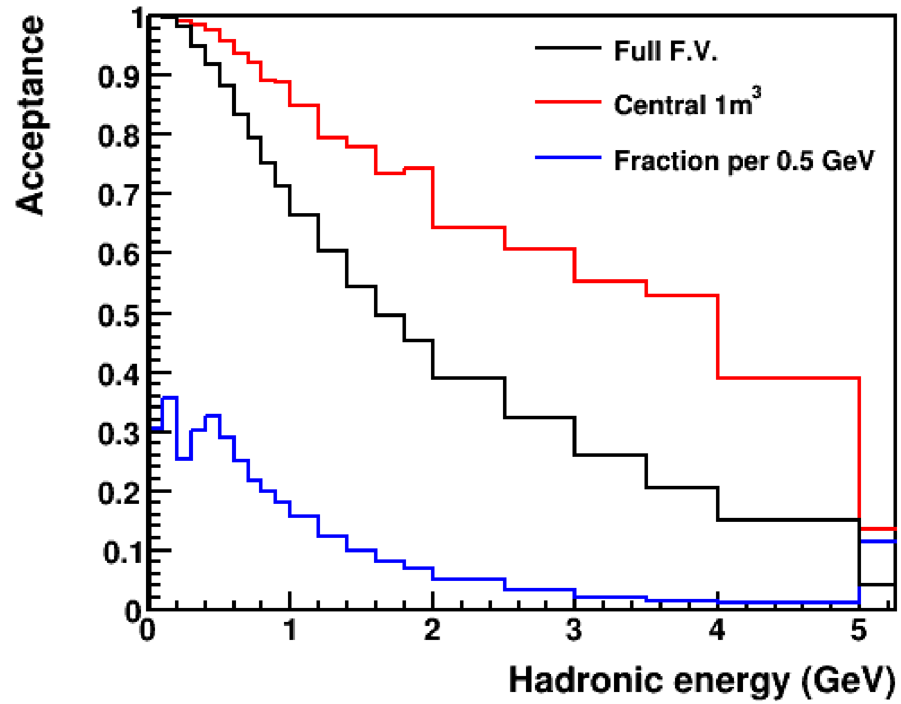
Chris Marshall

Hadronic energy reconstruction and containment requirement



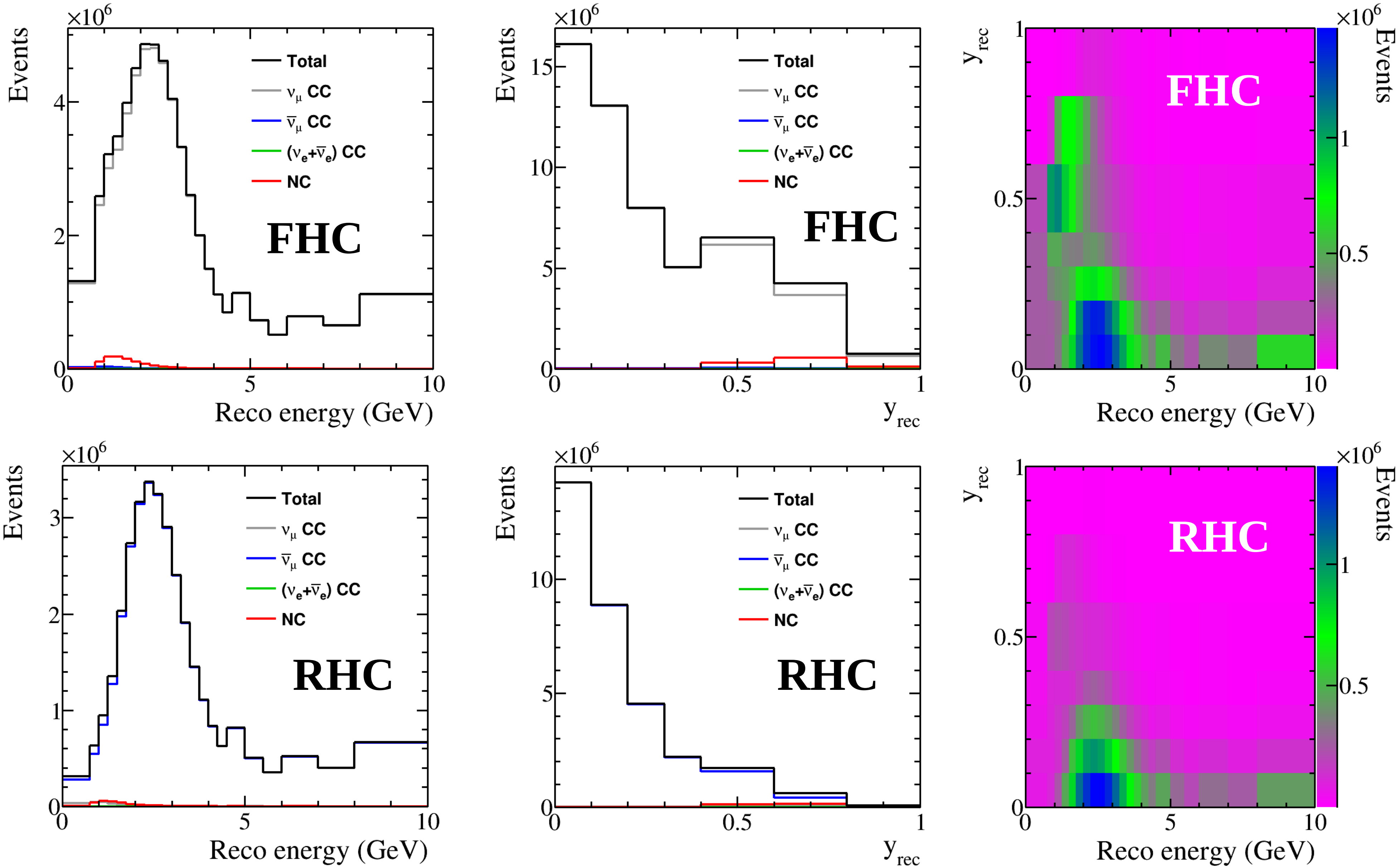
LAr ND event acceptance





- Left: CC acceptance vs. muon kinematics
- Right: Acceptance vs. hadronic energy – events with exiting hadrons are rejected, hence the lower efficiency at very high energy

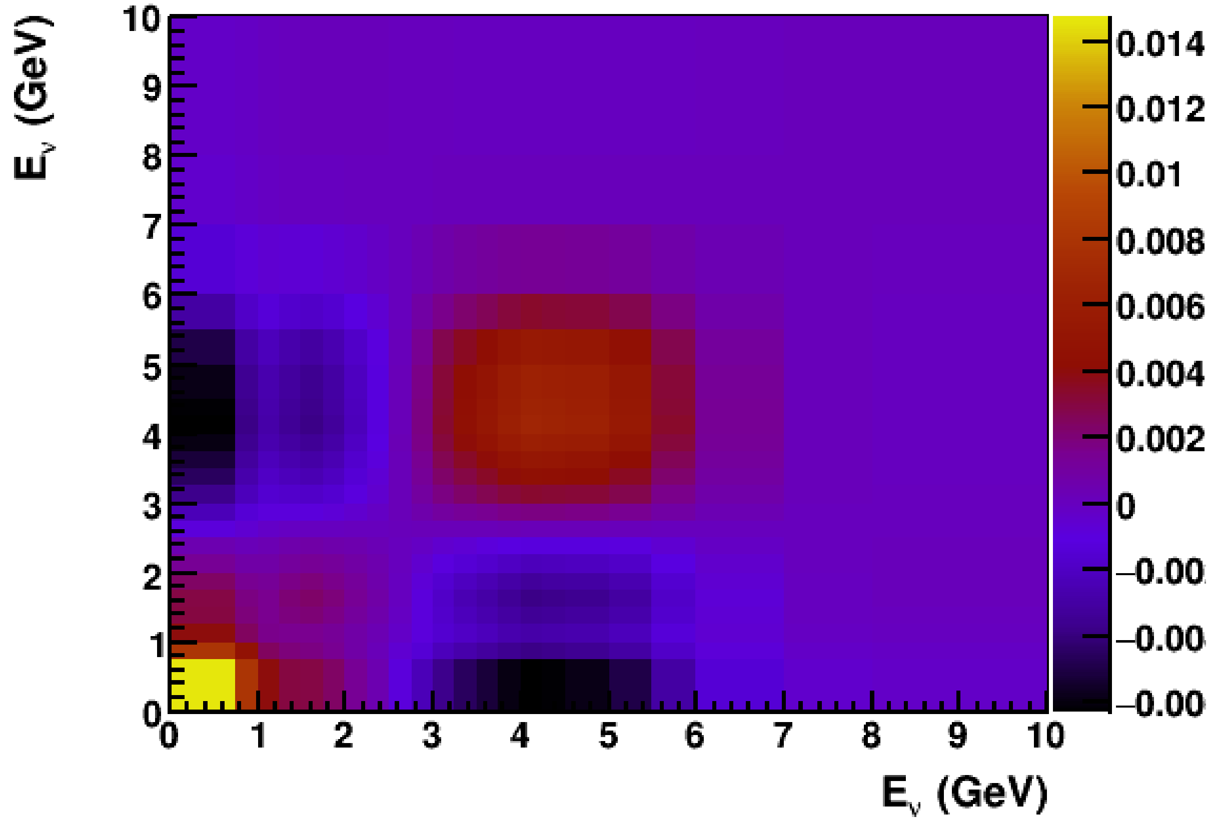
Selected ND analysis samples



Detector systematic uncertainties

- DUNE ND has O(100M) events → statistical uncertainties are negligible
- Critically important to have realistic systematics, despite not having a realistic simulation, reconstruction, or event selection
- Essentially impossible to implement detector uncertainties as nuisance parameters, because they will simply be determined in the fit

ND uncertainty implementation in covariance matrix

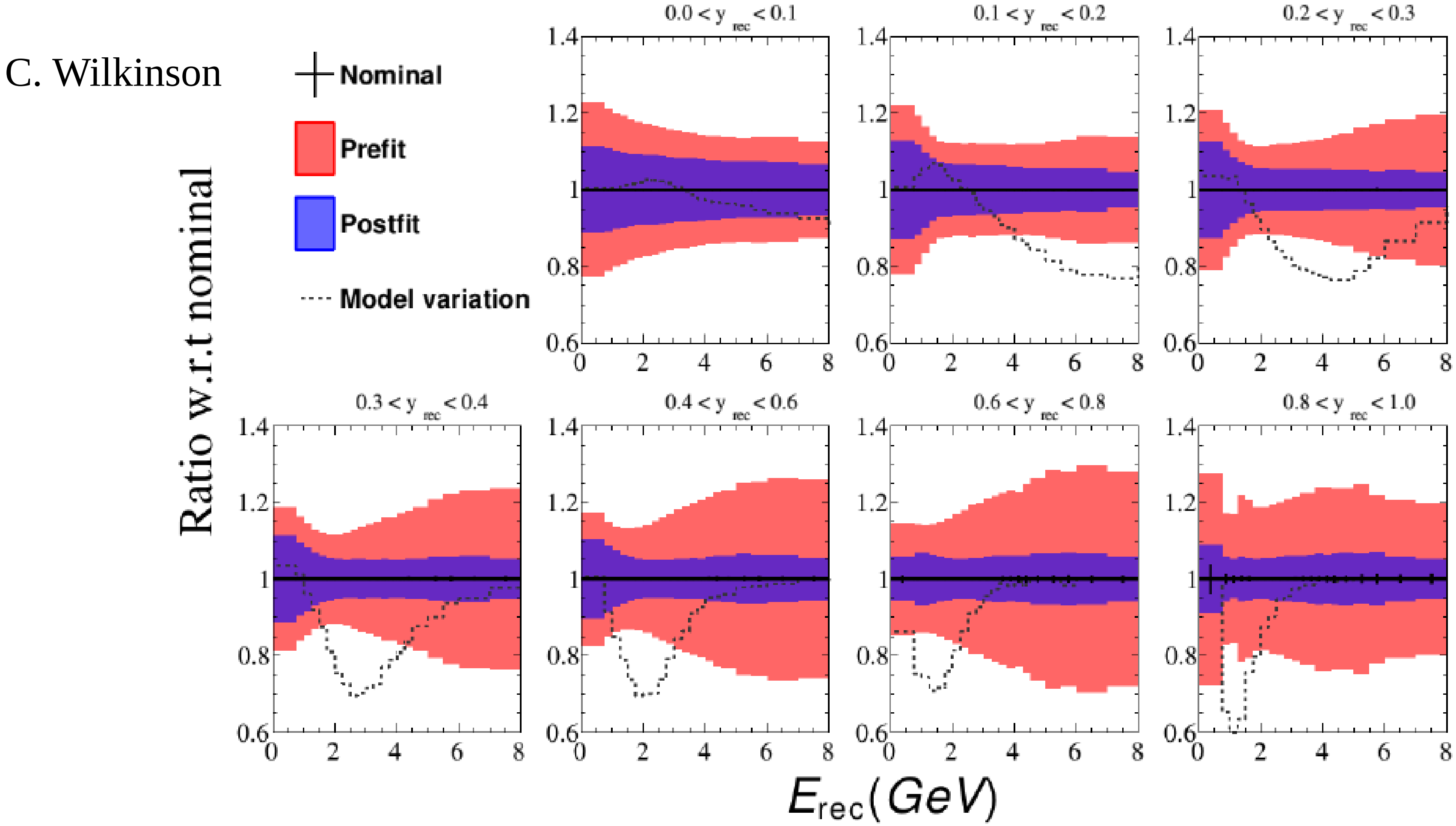


- Construct a covariance matrix in (E_ν, y) from the ND uncertainties with many universes approach
- Equivalent to including nuisance parameters in fit, but prevents these parameters from being further constrained by the ND data

Drawbacks

- Uses a single ND sample – not practical to directly implement dozens of possible selected samples in LAr, GAr, 3DST
- With covariance matrix you lose access to parameter constraints – difficult to show how the ND is constraining uncertainties
- Implicitly assumes that interaction and detector models are correct and describe the data, up to the included uncertainties
- Uncertainties only impact FD when there is degeneracy – i.e. two parameters that have the same effect on the ND – which never happens when you include enough bins with no statistical uncertainty

Example: MK single pion

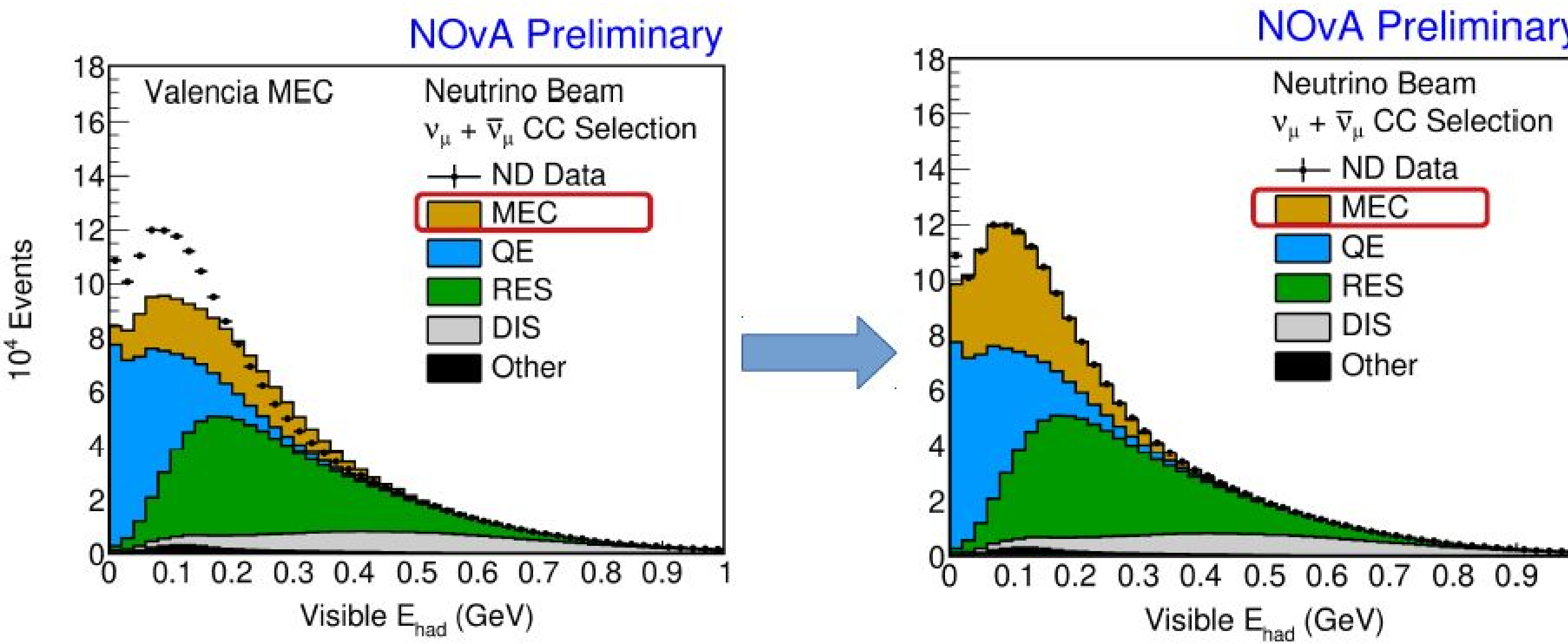


- Easy to see why this on/off dial (MK SPP reweight) is simply resolved by the ND... it simply knows whether it's on or off.

ND TDR analysis goals

- For ND TDR, we want to demonstrate how the various components of the ND each constrain specific uncertainties
- To do this, we need to add additional analysis samples, including selected events in the HPgTPC and 3DST
- Covariance matrix for detector uncertainties does not scale well
- Re-implementing detector parameters in the fit will cause them to be overconstrained → requires many, many more parameters → much longer fit time (already took ~37M CPU hours)

How it works in experiments



- ND data will **not** be described by our model
- We will modify our model to describe the ND data in many different projections, and add systematic uncertainties for the many different ways this can be done

Basic idea of mock data studies

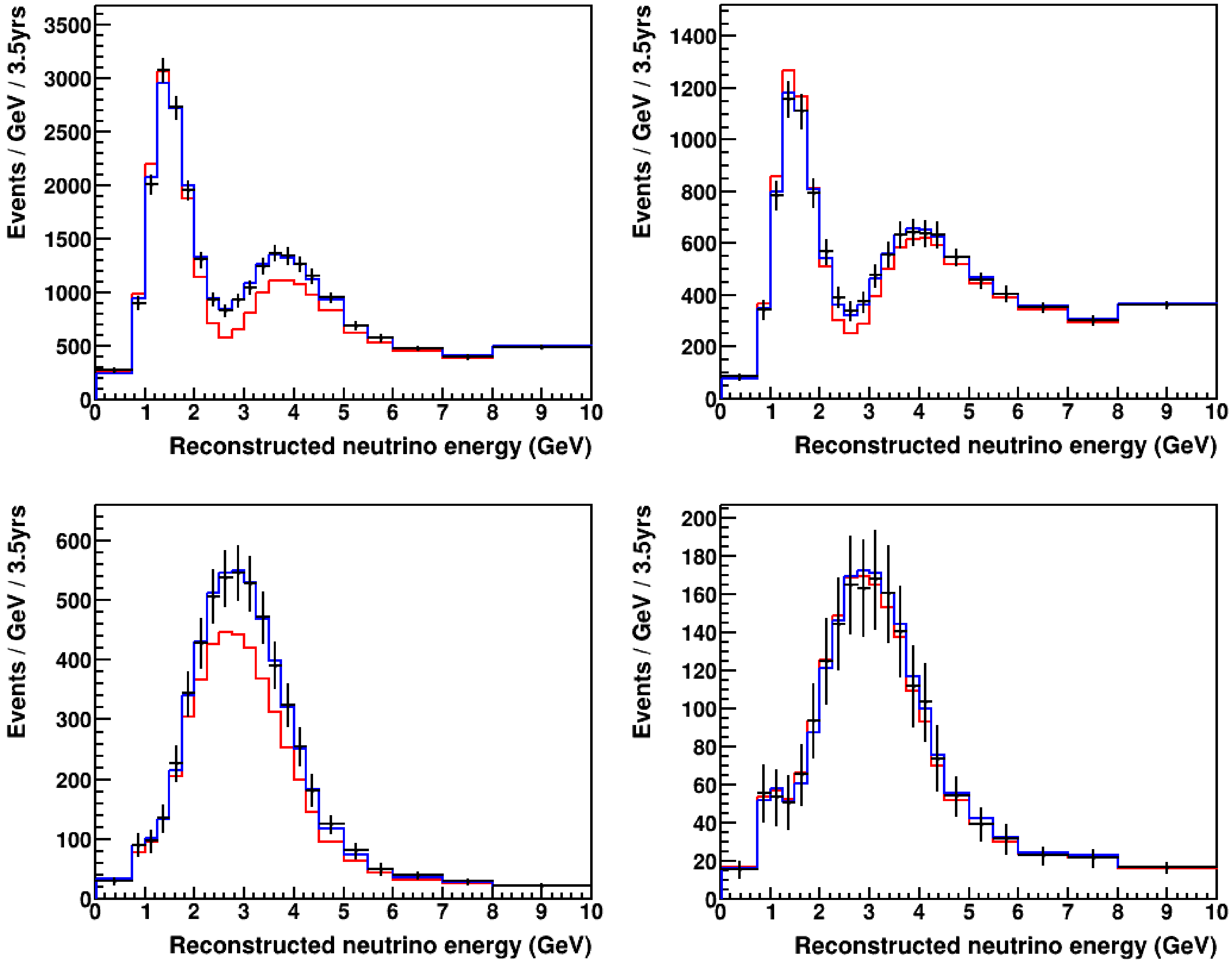
- Use an alternate model (i.e. not constructed by varying uncertainty parameters of reference model) to generate the "data" for ND and FD
- Determine the impact on oscillation parameters by fitting with FD only (or possibly FD + a limited ND)
- Demonstrate how a particular near detector sample clearly discriminates between the alternate and reference model

Example: NuWro mock data

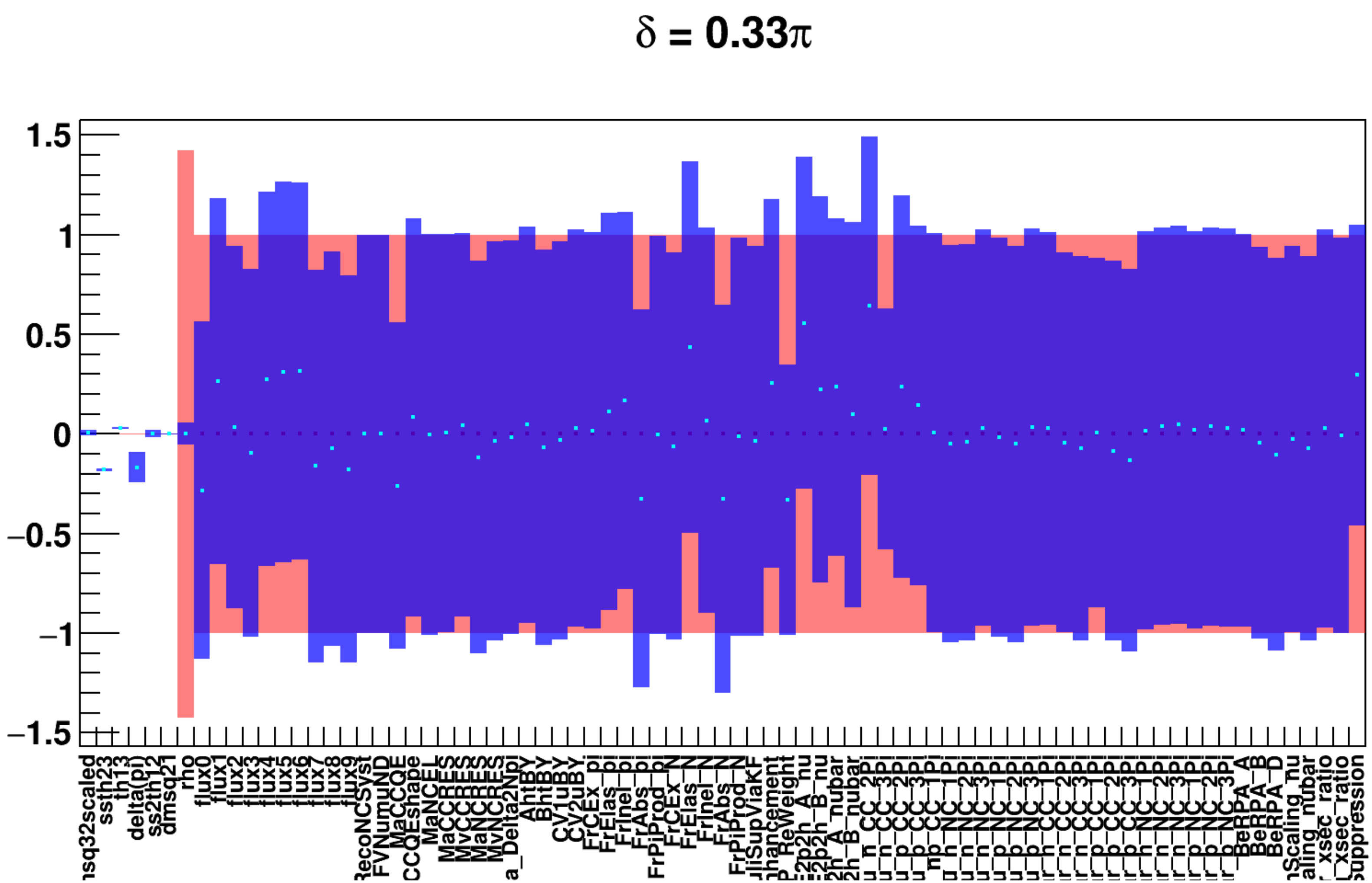
- Use Cris Vilela's reweighting tool to produce a mock data sample based on NuWro
- Fit with FD only, to show bias that could be expected in an experiment without a near detector
- Fit with ND+FD to show that reference model is very strongly disfavored

FD-only fits

- FD-only we get very good fit, with $\chi^2 \sim 10$
- No evidence of any problems with model

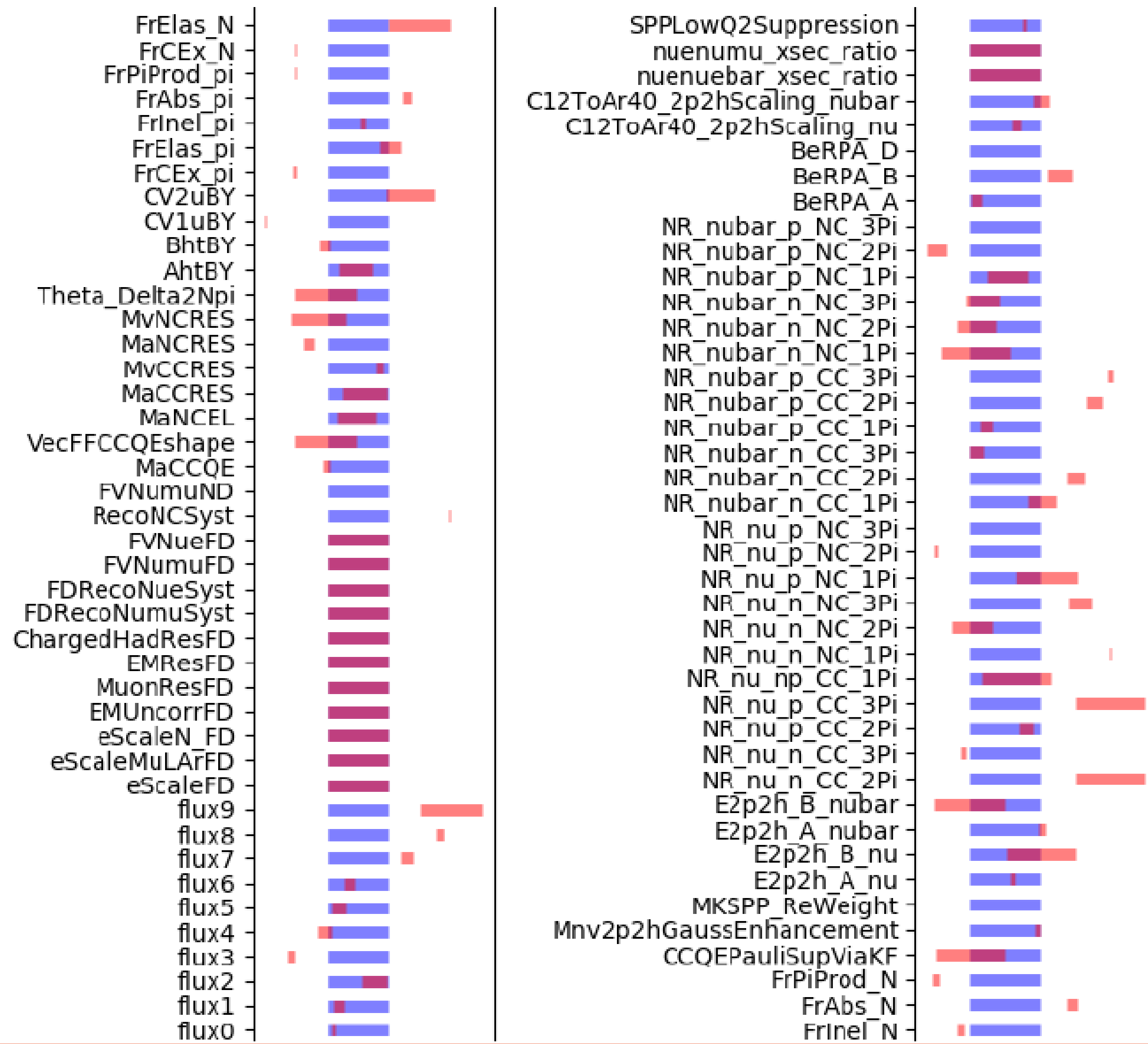


FD-only nuisance parameter post-fits are < 0.5σ of pre-fit values

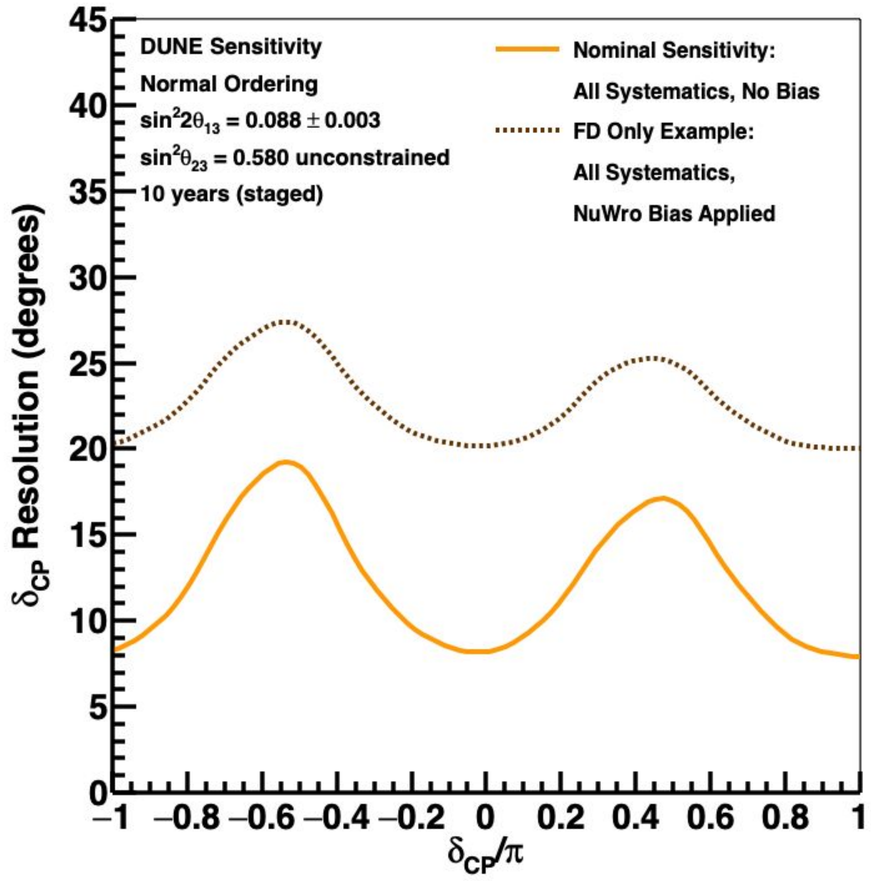


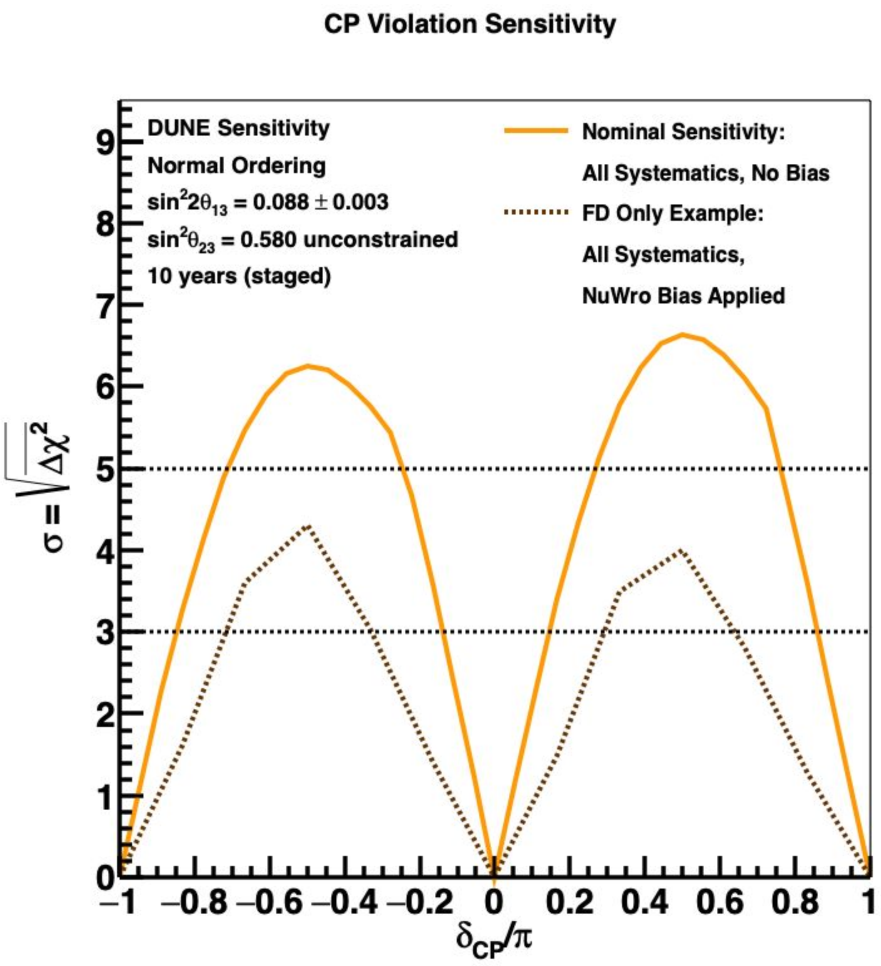
ND+FD fit χ^2 = 10879.2

- Post-fit parameter uncertainties are shown as red bands
- Parameters get pulled way outside their pre-fit ranges, with tiny constraints
- Fit to ND data is terrible – we would definitely know there is a problem, although we do not yet show how we would fix it



Sensitivities with bias applied





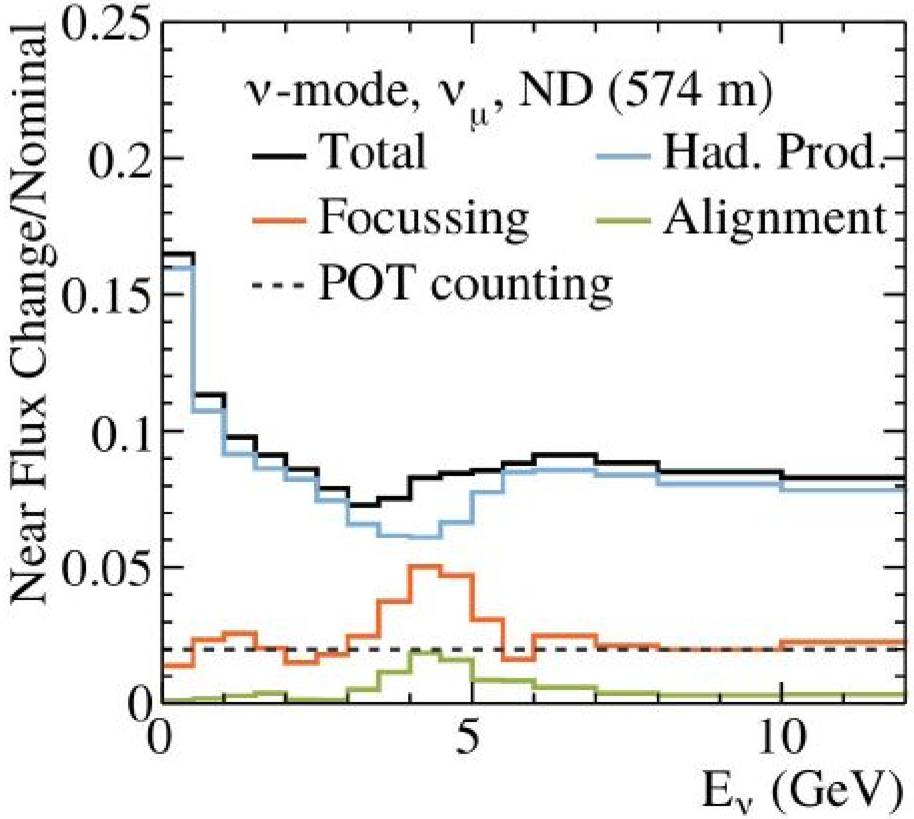
Creating mock data samples

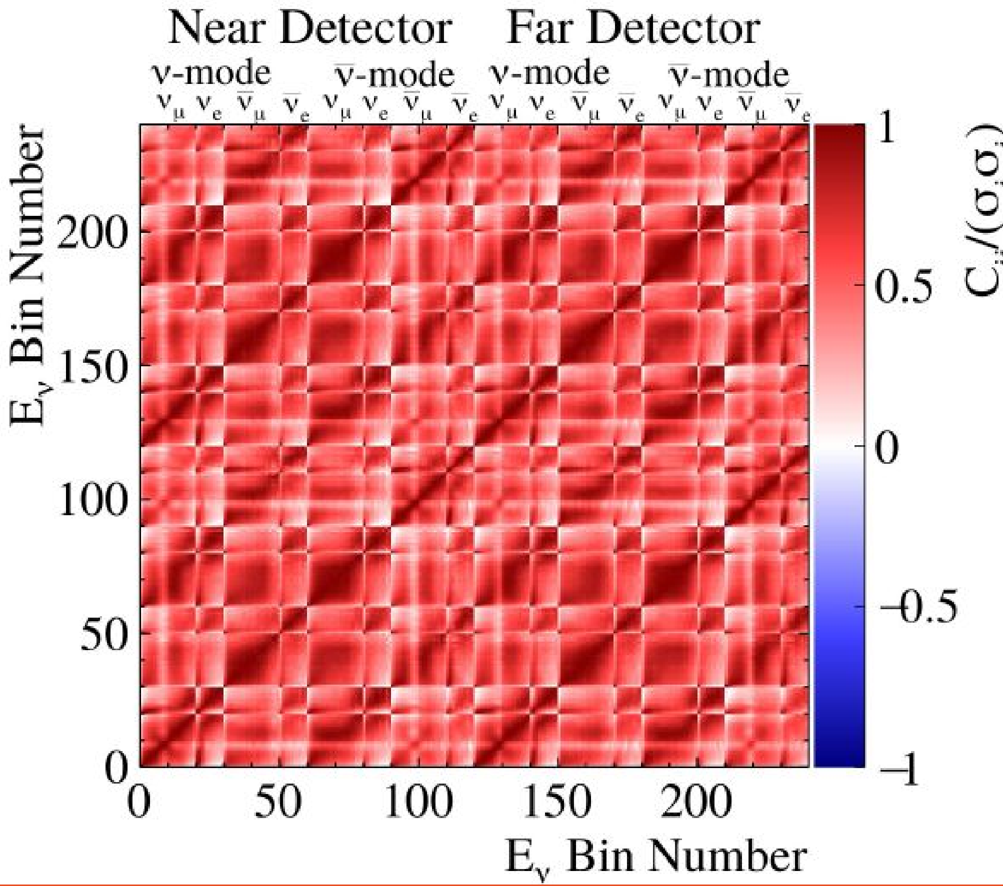
- Historically it is difficult to produce alternate samples
- Must re-run entire simulation chain, pseudo-reconstruction, etc. to produce analyzable files
- Cris Vilela's tool makes this process very straightforward: see next talk

Backups

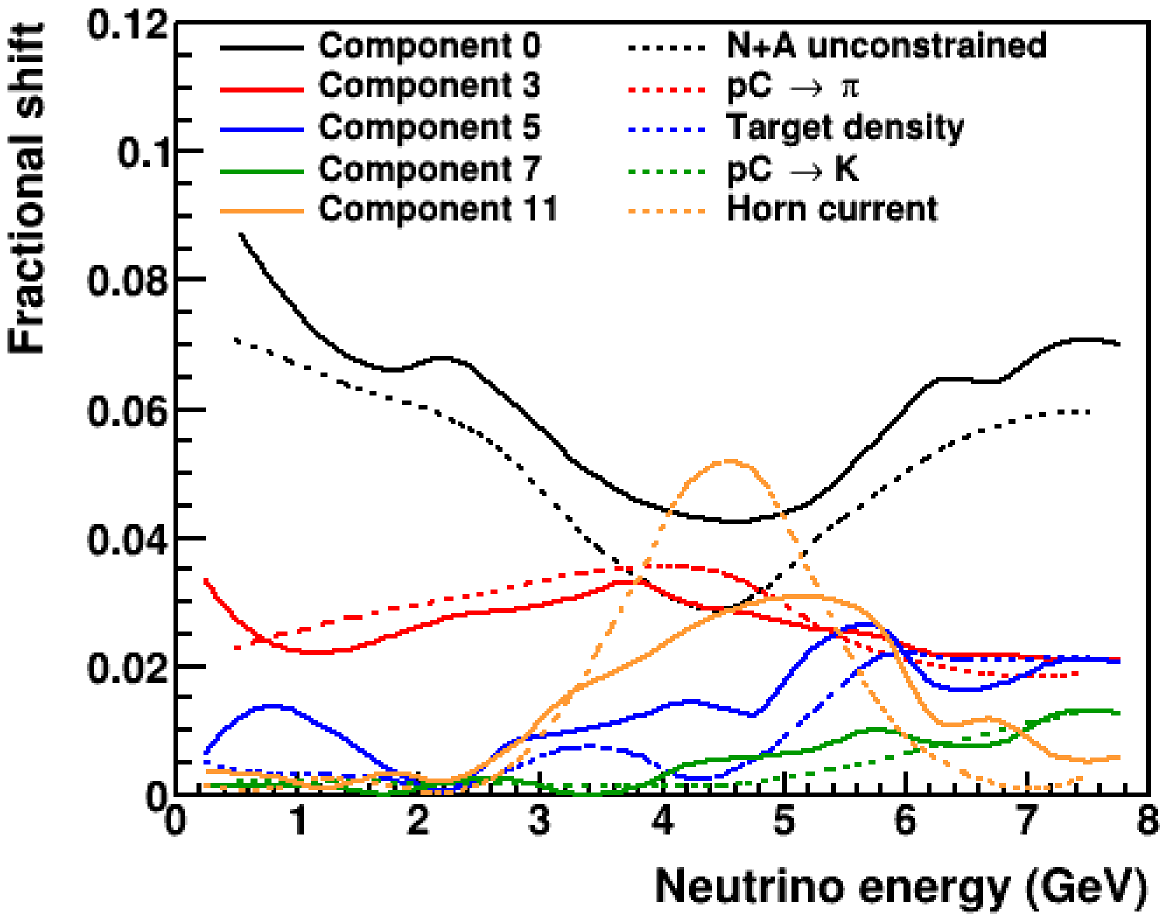
Flux uncertainties

- Flux uncertainties due to hadron production, beam focusing, and alignment are evaluated, including strong correlations between bins, beam modes, neutrino flavors





Principal component analysis is used to improve performance



- The largest HP & focusing uncertainties show up as principal components of the full covariance matrix
- Validates that our mathematical trick to diagonalize the uncertainty captures the same physics as varying individual parameters

Cross section uncertainties

MaCCQE
VecFFCCQEshape
CCQEPauliSupViaKF
MaNCEL
MaCCRES
MvCCRES
MaNCRES
MvNCRES
Theta_Delta2Npi
AhtBY
BhtBY
CV1uBY
CV2uBY
FrCEx_pi
FrElas_pi
FrInel_pi
FrAbs_pi
FrPiProd_pi
FrCEx_N
FrElas_N
FrInel_N
FrAbs_N
FrPiProd_N

Mnv2p2hGaussEnhancement
MKSPP_ReWeight
E2p2h_A_nu
E2p2h_B_nu
E2p2h_A_nubar
E2p2h_B_nubar
BeRPA_A
BeRPA_B
BeRPA_D
C12ToAr40_2p2hScaling_nu
C12ToAr40_2p2hScaling_nubar
nuenuebar_xsec_ratio
nuenumu_xsec_ratio
SPPLowQ2Suppression

NR_nu_n_CC_2Pi
NR_nu_n_CC_3Pi
NR_nu_p_CC_2Pi
NR_nu_p_CC_3Pi
NR_nu_np_CC_1Pi
NR_nu_n_NC_1Pi
NR_nu_n_NC_2Pi
NR_nu_n_NC_3Pi
NR_nu_p_NC_1Pi
NR_nu_p_NC_2Pi
NR_nu_p_NC_3Pi
NR_nubar_n_CC_1Pi
NR_nubar_n_CC_2Pi
NR_nubar_n_CC_3Pi
NR_nubar_p_CC_1Pi
NR_nubar_p_CC_2Pi
NR_nubar_p_CC_3Pi
NR_nubar_n_NC_1Pi
NR_nubar_n_NC_2Pi
NR_nubar_n_NC_3Pi
NR_nubar_p_NC_1Pi
NR_nubar_p_NC_2Pi
NR_nubar_p_NC_3Pi

GENIE ReWeight

MaCCQE
VecFFCCQEshape
CCQEPauliSupViaKF
MaNCEL
MaCCRES
MvCCRES
MaNCRES
MvNCRES
Theta_Delta2Npi
AhtBY
BhtBY
CV1uBY
CV2uBY
FrCEx_pi
FrElas_pi
FrInel_pi
FrAbs_pi
FrPiProd_pi
FrCEx_N
FrElas_N
FrInel_N
FrAbs_N
FrPiProd_N

GENIE reweight parameters affecting
CC quasi-elastic
CC resonance production
CC deep inelastic scattering
Final-state interactions
Neutral currents

DUNEint not covered in GENIE

Additional parameters:

CC QE

CC Resonance

2p2h

Scaling C → Ar

ν_e/ν_μ or $\nu_e/\bar{\nu}_e$

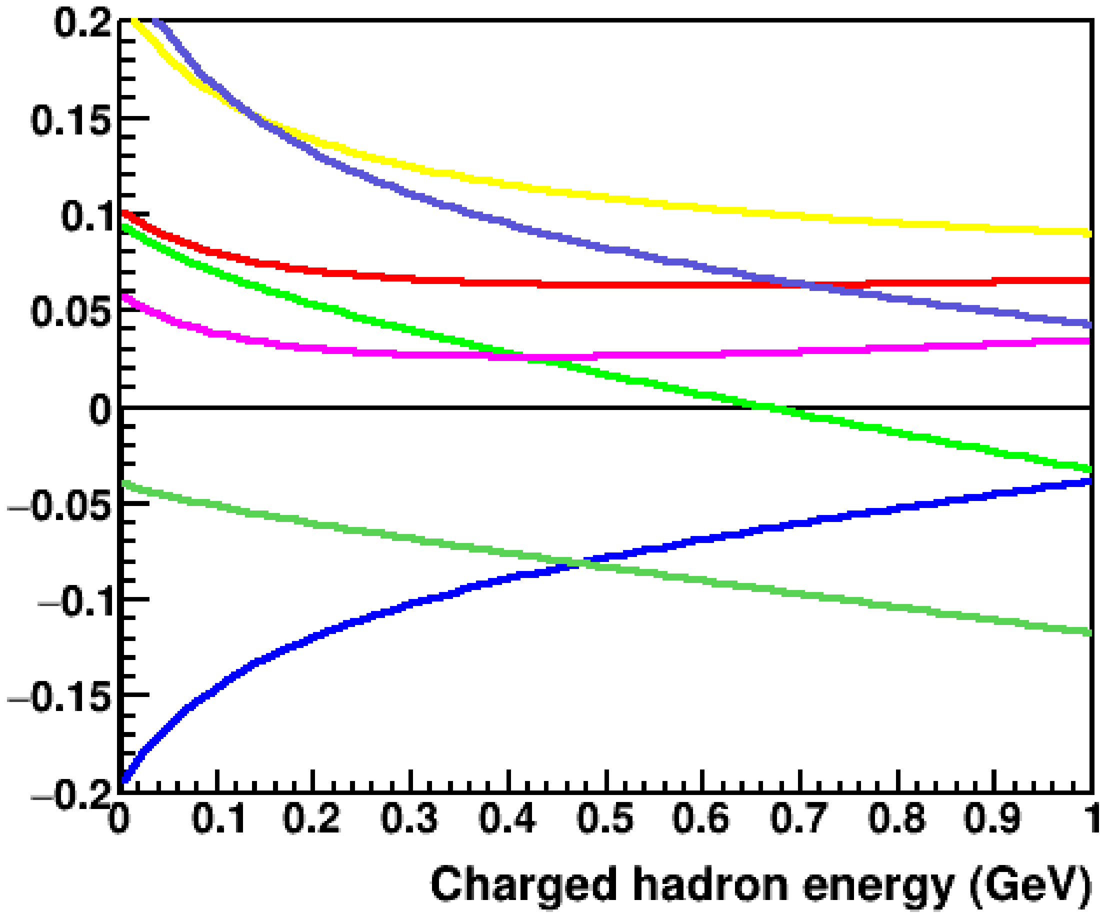
Mnv2p2hGaussEnhancement
MKSPP_ReWeight
E2p2h_A_nu
E2p2h_B_nu
E2p2h_A_nubar
E2p2h_B_nubar
BeRPA_A
BeRPA_B
BeRPA_D
C12ToAr40_2p2hScaling_nu
C12ToAr40_2p2hScaling_nubar
nuenuebar_xsec_ratio
nuenumu_xsec_ratio
SPPLowQ2Suppression

DUNEint not covered in GENIE

Additional parameters affecting non-resonant pion production

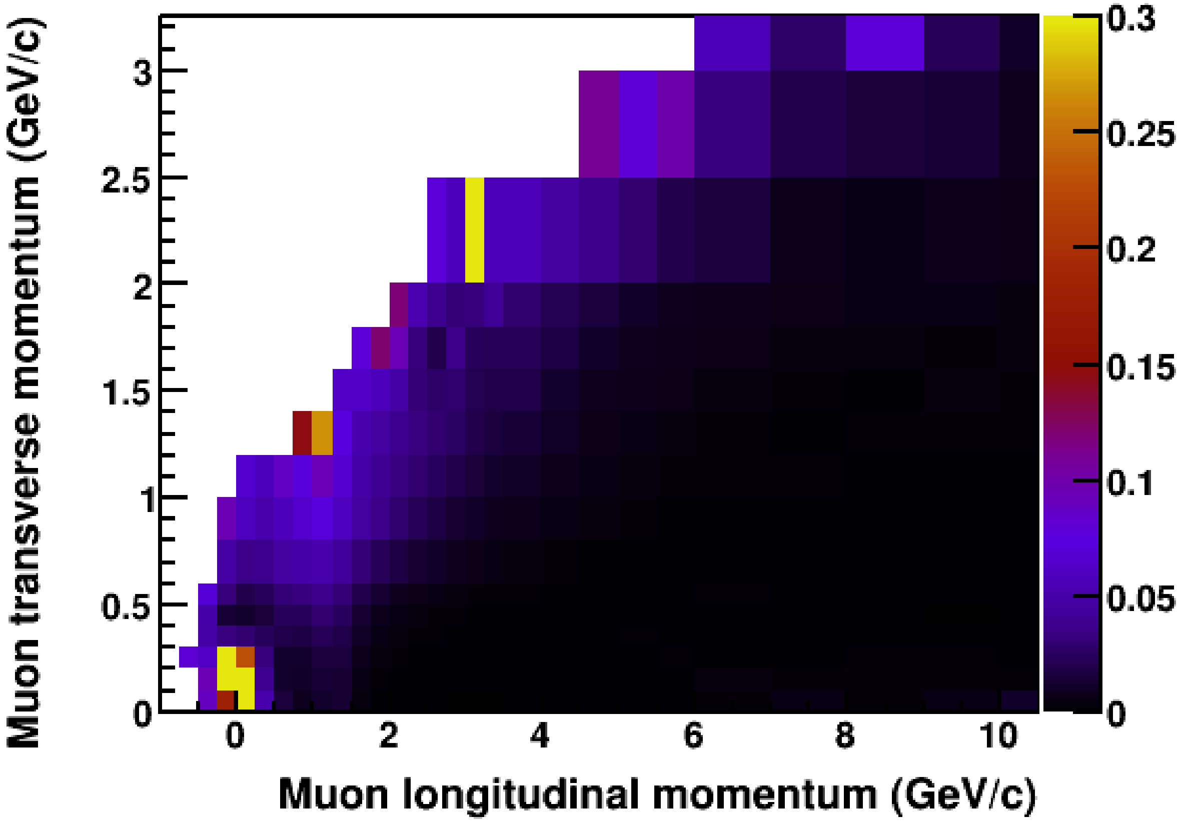
NR_nu_n_CC_2Pi
NR_nu_n_CC_3Pi
NR_nu_p_CC_2Pi
NR_nu_p_CC_3Pi
NR_nu_np_CC_1Pi
NR_nu_n_NC_1Pi
NR_nu_n_NC_2Pi
NR_nu_n_NC_3Pi
NR_nu_p_NC_1Pi
NR_nu_p_NC_2Pi
NR_nu_p_NC_3Pi
NR_nubar_n_CC_1Pi
NR_nubar_n_CC_2Pi
NR_nubar_n_CC_3Pi
NR_nubar_p_CC_1Pi
NR_nubar_p_CC_2Pi
NR_nubar_p_CC_3Pi
NR_nubar_n_NC_1Pi
NR_nubar_n_NC_2Pi
NR_nubar_n_NC_3Pi
NR_nubar_p_NC_1Pi
NR_nubar_p_NC_2Pi
NR_nubar_p_NC_3Pi

Example: charged hadron response



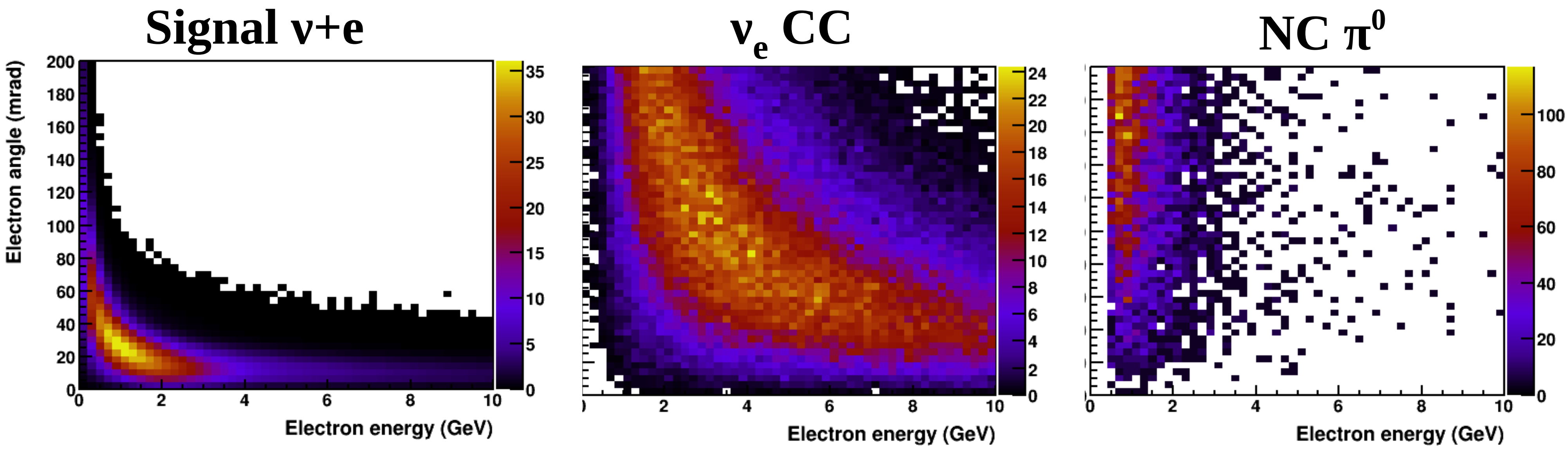
- Each curve represents the energy response bias in a particular universe, where the parameters have been chosen randomly consistent with the energy-dependent uncertainty

ND CC ν_μ acceptance fractional uncertainty



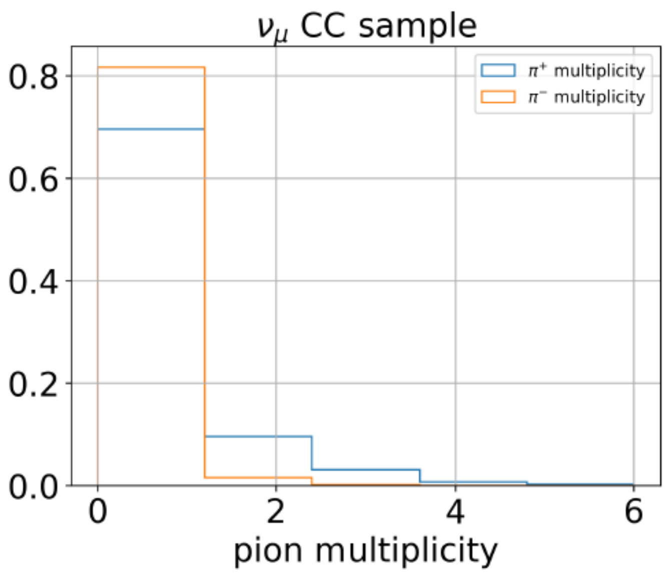
- CC events are rejected when
  - Muon is reconstructed as $\pi^{\pm}$ (low energy)
  - Muon exits sides
  - Muon exits downstream but does not enter gas TPC
- Acceptance is sensitive to detector modeling in phase space where muon acceptance is rapidly changing
- Uncertainty is evaluated as a function of muon momentum in transverse and neutrino direction (equivalently, energy and angle)

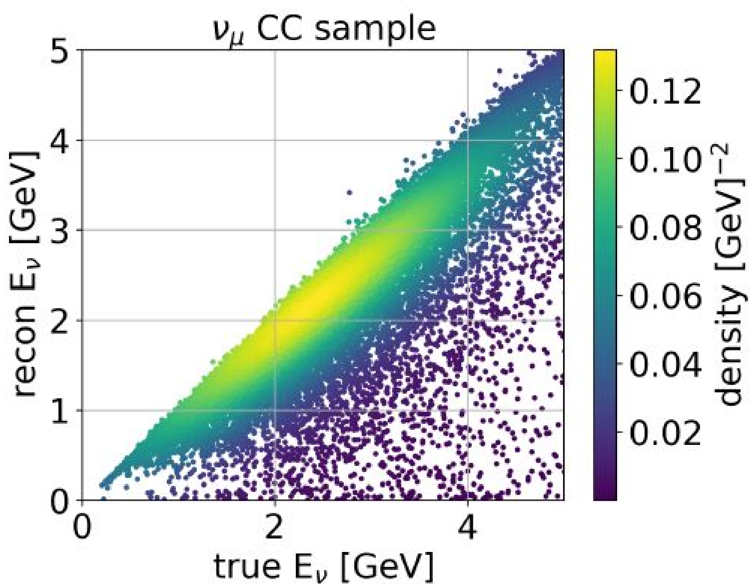
Additional LAr sample: ν+e scattering



- Pure EW process with known cross section → sensitive to flux only
- Signal is subject to kinematic constraint $E_e\theta_e^2 < 2m_e$
- Dominant background is ν_e CC at low Q^2
- Signal and background samples are ready, but have yet to be included in fit

Additional sample: Gas TPC





- Leverage low threshold, excellent PID of gas TPC, with very different detector systematics
- Binned in reconstructed E_ν separately for CC0π, 1π, >2π
- Complements LAr TPC by constraining some cross section parameters that are hard to access with dense LAr